Ugi multicomponent reaction to prepare peptide-peptoid hybrid structures with diverse chemical functionalities

Manuel Hartweg,1 Charlotte J. C. Edwards-Gayle,2,3 Elham Radvar,1 Dominic Collis,1 Mehedi Reza,4 Michael Kaupp,1 Jan Steinkönig,5 Janne Ruokolainen,5 Robert Rambo,3 Christopher Barner-Kowollik,5 Ian W. Hamley,4 Helena S. Azevedo1 and C. Remzi Becer1*

Monodisperse sequenced peptides and peptoids present unique nano-structures based on their self-assembled secondary and tertiary structures. However, the generation of peptide and peptoid hybrid oligomers in a sequence-defined manner via Ugi multicomponent reaction has not yet been studied. Herein, we report a synthetic strategy that enables both the modification of peptides as well as the generation of sequence-defined peptide-peptoid hybrid structures. Our synthetic methodology rests on the fusion of solid phase peptide synthesis with Ugi multicomponent reactions. We evidence that a diversity of chemical functionalities can be inserted into peptides or used in the design of peptide-peptoid hybrids exploiting a wide functional array including amines, carboxylic acids, hydrocarbons, carbohydrates as well as polymers, introducing a sequence-defined synthetic platform technology for precise peptoid hybrids.

Introduction

Nature has created sequence defined peptides to carry out specific functions, such as molecular recognition and biocatalysis, in the course of evolution. Today, even though synthetic macromolecules are widely used for a large number of applications, achieving an absolute control over their primary structure still remains as a key challenge for chemists.1-3 Therefore, our fundamental understanding of the role of monomer sequence and its influence on the secondary and tertiary structures of peptides and proteins has to be enhanced dramatically.

Amongst a range of methods for the synthesis of such sequence-defined macromolecules, Merrifield solid phase peptide synthesis (SPPS) method has been the most successful.4-6,7-17 Solid-phase protocols typically use excess of reagents and require two steps, which are the coupling and deprotection steps, for incorporation of each repeat unit.

Besides, extremely efficient reactions are required to ensure monodisperse formation of peptide chains in high yields.18,19 Complementary to the synthesis of peptides, Zuckermann et al. reported the synthesis of peptoids via solid-phase submonomer synthesis. Polypeptoids are accepted as peptidomimetic polymers20,22 with a substituent on the amide nitrogen, which significantly influences the formation of higher order structures due to suppression of a hydrogen bonding.23,24 Moreover, the N-substitution eliminates the chirality of the more flexible peptoid backbone,25 which transforms peptoids to a generally side-chain dominated system, dissimilar to peptides in which conformation and packing are dictated primarily by the chemical structure of the backbone. However, compared to peptides, engineered peptoids display a range of interesting properties, such as protein-mimetic secondary and tertiary structures.26-29 They have been demonstrated to form stable nanosheets in vivo,30,31 as well as to improve tissue accumulation for reduced excretion rates.32,33 Further successful examples for the peptoid synthesis can be listed as solid-phase synthesis via Fmoc-strategy,34 bromoacetyl-bromide method in solution-phase,35 N-heterocyclic carbene mediated ring-opening polymerization of N-substituted N-carboxy-anhydrides36,37,38, and Ugi 4-component reaction (Ugi-4CR).39 However, up to date, only very little research was carried out on peptide-peptoid hybrids. For example, a few studies have reported the synthesis of sequence defined hybrids with various biological applications.40-44 Also the synthesis of peptide-peptoid block conjugates,42,45,46 using native chemical ligation (NCL),47 fragment condensation,47 and copper-catalyzed azide-alkyne [3 + 2] cycloaddition48 have been reported. In the present study, Ugi-4CR was utilized in order to generate highly
functional and sterically demanding peptoid units. Recently, the Ugi-4CR was already applied for the sequence defined polymers, the functionalization of peptide arrays\textsuperscript{49} synthesis of polypeptoids from $\gamma$, $\delta$, $\epsilon$-amino acids,\textsuperscript{50} and alternating peptide-peptoid copolymers from\textsuperscript{51} Ugi-4CR chemistry generally displays the high efficiency,\textsuperscript{52-55} compatibility with amino acid chemistry,\textsuperscript{56} while allowing incorporation of a range of chemical diversity into peptoid side chains.\textsuperscript{57} The method benefits from its simplicity and remarkably high diversity for modification of peptides or synthesis of novel peptoid and hybrid structures.

Herein, we report an iterative synthesis method towards highly functional, sequence-defined, and monodisperse peptides, peptoids, and peptide-peptoid hybrids (Fig. 1). In an extended solid-phase protocol, the methodology fuses SPPS and Ugi-4CR chemistry for the incorporation of peptide and peptoid units, respectively. Furthermore, we investigate the self-assembly behavior and defragmentation pattern of selected compounds to demonstrate the diversity of these hybrid structures. Moreover, we propose a nomenclature for peptoid structures generated by the means of Ugi-4CR, as well as for peptoid Ugi-peptide hybrids.

![Figure 1. Schematic representation of (a) conventional solid-phase peptide synthesis (SPPS), (b) Solid-phase peptoid synthesis via Ugi reaction, and (c) the iterative process for the synthesis of peptide-peptoid hybrids via combination of SPPS and Ugi-SPPS (AA = amino acid; $R_{AA} = R_{aldehyde}$; $R_{IC} = R_{isocyanide}$).](image)

RESULTS AND DISCUSSION

The preliminary considerations indicated that the distinct reaction protocols of conventional SPPS and Ugi-SPPS are suitable for iterative, random, or programmed sequence-defined solid phase synthesis. Therefore, a nomenclature was suggested (ESI 1), indicating the whether a unit is of peptide or peptoid nature, and which $R$-group the peptide ($R_{AA}$) or two independent $R$-groups the Ugi-peptoids ($R_{AA}$ and $R_{IC}$) are carrying. However, for the Ugi-SPPS process, the combinations of several amino acids (i.e. glycine, phenylalanine and lysine), aldehydes (i.e. butyraldehyde, undecanal, undecenal, glucose aldehyde, and PEG-aldehyde), and isocyanides (cyclohexyl, and 2-morpholinoethyl isocyanide) were investigated. An Fmoc-Rink-amide coupling strategy was employed throughout the process for both coupling methodologies (Fig. 1). The process typically breaks down into six steps: (i) automated deprotection of the resin, (ii) automated Fmoc-solid phase peptide synthesis (Fmoc SPPS) of the anchors, (iii) manual Fmoc-SPPS (iv) manual Ugi-Fmoc-solid phase peptoid synthesis (Ugi-Fmoc-SPPS), (v) cleavage and (vi) purification. In order to validate the successful combination of SPPS and Ugi-SPPS, compounds 1-6 (Fig. 2) were prepared. Thus, peptide anchors (H$_2$N-FGF-CONH$_2$ in 1, 2, 7, H$_2$N-KKK-CONH$_2$ in 3-6, 8-12) were prepared via automated SPPS on Rink-amide resin. Anchors were then deprotected to generate the free amines for incorporation of subsequent peptide or peptoid units. Afterwards, in contrast to conventional SPPS, where the amine moieties are reacted with coupling reagents and an amino acid, amines were reacted with an excess of aldehyde (4.00 eq) in 1:3 MeOH/CHCl$_3$. It is important to note that the use of both MeOH and CHCl$_3$ was crucial to achieve full and selective
conversion of the terminal amine into a peptoid sequence. CHCl₃ is essential to ensure sufficient swelling of the resin, whereas MeOH is required to promote Ugi reaction. After complete imine formation, the respective Fmoc-protected amino acid (5.00 eq) and isocyanide (6.00 eq) were added to react for a further 2-6 hours. Terminal free amine has been formed as a result of the subsequent deprotection step in 1:4 Pyridine/DMF. Successive coupling steps were performed accordingly to obtain the desired sequence of the final peptoid or peptide/peptoid hybrids. Following the last coupling reaction or the final deprotection step, oligomers were cleaved off the solid support with 1:1 TFA/DCM mixture and crude mixtures were analyzed. Electrospray ionization mass spectrometer (ESI-MS) or matrix assisted laser desorption ionization time of flight mass spectrometer (MALDI-ToF/MS) were used to investigate the conversion and chemoselectivity of oligomers (Fig. 3). Furthermore, reverse phase high performance liquid chromatography (RP-HPLC) or size exclusion chromatography (SEC) was conducted to confirm the purity of monodisperse oligomers. SEC analysis of 1 and 2 have displayed narrow peaks and indicate the formation of the desired sequences (Fig. S6-S7).

Subsequently, to obtain more polar hybrids, H₂N-KKKK-CONH₂ was used as an anchor. The reactions with undecenal or butyraldehyde, Fmoc-Phe-OH, and cyclohexyl isocyanide, as

Figure 2. Chemical structures (1-12) of the peptide/peptoid hybrid structures synthesized in this study. Green and blue units represent peptide and peptoid units, respectively. The letter codes for the structures are displayed according to ESI 1.
Figure 3. Mass spectrometric analysis of peptide-peptoid hybrids (1-9) and oligopeptoid (10-12) arrays. Experimental and theoretical (in parenthesis) m/z values that provided ESI-MS (1-5, 7-10) or MALDI-ToF/MS (6, 11, 12) are shown. In all cases, MS analyses are performed directly on crude products that are cleaved off the resin to demonstrate the efficiency of this iterative process.

well as successive SPPS with Fmoc-Phe-OH, yielded oligomeric 3 and 4 at high purity based on RP-HPLC traces (Fig. 3). Compound 3 has two main isomers, one from cis and one from the trans, of the bis-amide formed as a result of Ugi-4CR. Furthermore, ESI-MS analysis was carried out for 3 and 4, where mainly singly and doubly charged cations were observed (Fig. 3). Subsequently, the addition of a glucose aldehyde resulted in a chemoselective formation of glycopeptoid-peptide hybrid 5 that was confirmed by RP-HPLC and ESI-MS (Fig. S14 and Fig. 3). An acetyl-protected sugar-derived aldehyde was synthesized via base catalyzed thiol-Michael addition using tetra-acetylated thioglucose and acrolein (ESI). Additionally, Ugi-4CR was used as a ligation tool in the synthesis of the hybrid 6. Therefore, PEG-aldehyde was synthesized by Swern oxidation reaction (ESI).58 Thus, PEG<sub>500</sub> was used as a starting material and ligated to the anchor compound 6. The success of the reaction was determined by MALDI-ToF, whereby the PEG signals shift...
Figure 4. MS/MS analysis of peptide-peptoid hybrids. 

**a.** Top: Schematic presentation of compound 8 and its assignment into different fragmentation patterns. 
**Bottom:** List of unit, and dissociation patterns of main chain built by amino acids, as well as side chain of peptoid units of compound 8. The amino acids are abbreviated with their respective one letter code, the R-groups of the side chain with their chemical composition. 

**b.** MS/MS analysis of compound 8. 

**Blue trace:** isolated parent ion \([M + Na]^+\). 

**Green trace:** MS/MS at lower energy. 

**Gold trace:** MS/MS at higher energy.

This journal is © The Royal Society of Chemistry 20xx

J. Name., 2013, 00, 1-3 | 5
Polymer Chemistry

by $\Delta m/z = 917.2$ Da (Fig. S15), which corresponds exactly to the mass of peptoid/aldehyde.

To further expand the versatility of this synthetic process, several hybrids with different backbone architectures, such as alternating 7, and 8 and block 9 co-oligomers, were prepared. Starting from the same H$_2$N-FGF-CONH$_2$ peptide anchor, three peptoid sequences were attached with a phenylalanine peptide sequence in between the peptoids. The MS analysis of compound 7 (Fig. S17) demonstrated high purity even in the absence of any purification steps. Also, H$_2$N-KKKK-CONH$_2$ was designed as a more polar anchor and used in the synthesis of 8 and 9. ESI-MS (Fig. S19) shows the formation of the 9-mer displaying the singly, doubly and triply charged cations. Compound 9 is designed as a glycopeptoid containing three glycan units. Acetyl-protected glucose aldehyde was used in the synthesis and following to the deacetylation of glucose units, ESI-MS revealed signals for the doubly and triply charged cations (Fig. 3). RP-HPLC analysis of compound 9 also indicated the formation of an oligomer in high purity.

In the final set of hybrid structures, homopeptoids were studied to obtain highly functional oligopeptoids. Initially, 10 was prepared and analyzed as a trimer. MS analysis (Fig. 3) indicated a small amount of a side product, which was assigned as the trimer incorporating only two peptoid and one peptide sequence instead of the third peptoid unit. The main reason for this effect is ascribed as the steric hindrance as the reaction rate significantly suffers when bulky reagents are used in the Ugi reaction. The isotopic pattern of the main peak (Fig. S24) is in good agreement with the theoretical.

Figure 5. Cryo-TEM, CD and SAXS data for compounds 3, 5 and 8. a. Compound 3, 1 w% solution, 20 v% MeOH. b. Compound 5, 1 w% solution. c. Compound 8, 10 v% THF. d. CD spectra of compounds 3, 5, and 8. e. SAXS profiles scattering intensities of 1 w% solution of 3 in 20 v% MeOH and 80% H$_2$O (Green), 1 w% solution of 5 in 100% H$_2$O (black), and 1 w% solution of 8 in 10 v% THF and 90% H$_2$O (blue). Corresponding fits are shown as red lines. Intensity profiles for 3 and 5 are fitted to cluster models, data for 8 is fitted to a spherical shell micelle model.
Moreover, hexamer compound 11 using different starting compounds and heptamer compound 12 with identical units and molecular weights of approximately 3 kDa were synthesized. In the case of 11, the MALDI-ToF/MS analysis indicated that when 2-morpholinoethyl isocyanide was used alternatively, the reaction was not chemoselective and thus only one single peptide sequence was incorporated in the backbone (Fig. 3). MALDI-ToF/MS spectrum of 12 shows a series of peaks with a low intensity and \( \Delta m/z \) of one single Ugi-repetition unit was detected (Fig. 3). Nevertheless, the major peak was assigned to be the desired compound 12 and the isotopic pattern of 8 and 9 are found to be in good agreement.

Having demonstrated the versatility of the synthetic process to prepare sequence-defined peptide-peptoid hybrid structures, we then investigated their properties, namely their ability for information storage and self-assembly into nanostructures. Sequential deciphering of hybrid oligomers can be complicated. Due to the use of a multicomponent reaction, both peptide as well as peptoid units may be present within the hybrid. The selected analysis technique must be able to determine whether a sequence has a peptidic or peptoidic nature, and in case of a peptoid unit it should provide information regarding the composition of two incorporated R-groups derived from the Ugi-4CR. Although various analytical tools for the characterisation and sequencing of macromolecules have been developed in the past decades, biopolymers are usually decoded fast and reliably via tandem mass spectroscopy (MS/MS). However, no universal rules apply for the MS/MS dissociation patterns of non-natural polymers. For the specific fragmentation pathways, to decipher structural information of the peptide/peptoid hybrid macromolecules one has to refer to established polymer and peptide fragmentation assignments.

As illustrated (Fig. 4a), the composition of amino acids of peptide units can be assigned using their one letter code (i.e. K = lysine, F = phenylalanine). The nature of Ugi-peptoid units and their R-groups derived from the aldehyde and isocyanide can be assigned by the one letter code for the amino acid, as well as the abbreviations for the respective side chain units (i.e. nPr, Cy). In order to demonstrate the power of data storage on designed peptide-peptoid hybrid structures, compound 8 was analysed in detail via ESI-MS/MS. The parent ion (\([M + Na]^+\) (m/z = 2053.26 g·mol\(^{-1}\))) was isolated and activated with increasing collisional energy (Fig. 4b blue trace, SI Section 4). As expected, even at lower energies, a fragment (Z\(_9\)) was detected initially, which displays the cleavage of the urethane bond of the Fmoc-group (\( \Delta m/z = 222 \text{ g·mol}^{-1} \)). When the collisional energy was increased (Fig. 4b green and gold spectrum), further fragmentation patterns were observed. To our delight, we have not only observed fragmentation patterns from the main chain (C, M, N and Z), but also the specific side chain patterns could be identified. These outcomes provided structural information about both R-groups of the N-substitution of the peptoid units and thus revealing the location of every starting compound used in the respective Ugi-reaction. This was enabled by the bis-amide nature of the formed Ugi-scaffold, whereby the side chain amide moiety was fragmented on both sides of the amide bond (P, Q, O, as well as K and L). For instance, the fragmentation of the peptoid sequence situated at the eight sequence can be envisioned (N\(_8\) vs O\(_8\) vs Y\(_8\)). The fragment of the oligomer including sequence eight, N\(_8\) \( (m/z = 1685.09 \text{ g·mol}^{-1}) \) displays \( \Delta m/z = 127.12 \text{ g·mol}^{-1} \) compared to O\(_8\) \( (m/z = 1557.97 \text{ g·mol}^{-1}) \), which corresponds to CyNHC(H)=O. Additionally, Y\(_8\) \( (m/z = 1500.95 \text{ g·mol}^{-1}) \) provides information about the aldehyde derived R-group. A \( \Delta m/z = 57.02 \text{ g·mol}^{-1} \) value compared to Y\(_8\) belongs to CH\(_3\)(CH\(_2\))\(_2\)CH\(_3\) and displays the use of butyraldehyde (resulting in nPr as R-group) as a monomer in the Ugi-reaction. Similar observations were made for the sequence at location six and four. These observations were evidenced from \( \alpha \rightarrow \omega \) (N, O, P, Q and Y patterns) and \( \omega \rightarrow \alpha \) (C, K, L, M and C patterns) for each peptoid sequence, whereas the peptoid unit based fragments only revealed conventional peptide signals (C and Y patterns). For peptide and peptoid sequences, in addition X and Z as well as A and B patterns were observed, but neglected for clarity.

Secondary structures of 3, 5 and 8 were determined using circular dichroism (CD) spectroscopy. It was found that the structures were intrinsically disordered, identified through a negative peak around 200 nm (Fig. S29). The self-assembly properties of these three compounds were then investigated using cryogenic transmission electron microscopy (Cryo-TEM) and small-angle X-ray scattering (SAXS). Cryo-TEM images of 3 (Fig. 5a) revealed the presence of slightly larger globular structures with an average diameter of 14.95 nm. Compound 5 showed a small population of irregular sized aggregates or clusters (Fig. 5b). Finally, for compound 8 mixture of larger globular like structures and smaller micelle like assemblies with an average diameter of 3.125 nm are observed (Fig. 5c). SAXS was used to further probe the self-assembled structures of the hybrid oligomers, providing information on the structure and dimensions of the assemblies. Scattering intensities and form factor models are shown (Fig. S30). Scattering intensities of 3 and 5 were fitted based on a coexistence of clusters (described by a mass fractal form factor) and monomers (described by generalized Gaussian coil form factor). The models are in good agreement with cryo-TEM results, which showed a population of irregular aggregated structures. SAXS data for compound 8 fitted to a ‘spherical shell’ form factor model, corresponding to a core-shell micelle structure, also consistent with Cryo-TEM results. Parameters of fits are listed in Tables S2-S3.

Conclusions

In summary, two chemical synthesis techniques, i.e. solid-phase peptide synthesis (SPPS) and Ugi multicomponent reaction were combined to obtain peptide/peptoid hybrid structures with defined sequences and functionalities. Thus, a series of model peptide/peptoid hybrids were prepared. As evident from the various analysis techniques applied on the hybrids, both reactions proceeded at very high conversions and no further purification was required, demonstrating the power of the combined iterative approach and the diversity of the chemical groups that can be inserted in peptide sequences via Ugi reactions.
Acknowledgements

This work was partially supported by the European Commission Horizon2020 programme (EU-ITN EuroSequences Proposal No: 642083). C.R.B. also acknowledges EPSRC (EP/P009018/1) for the financial support. D.C. and H.S.A. thank the support from the EU-funded project “SuprHApolymers” (PCIG14-GA-2013-631871). I.W.H. thanks EPSRC (EP/L020599/1) for support and Diamond Light Source and the University of Reading for co-funding the Ph.D. studentship of C.E.G.. Diamond Light Source is thanked for the award of beamtime on beamline B21. C.B.K. acknowledges the Sonderforschungsbereich 1176 (project A3) funded by the Sonderforschungsbereich 1176 (project A3) funded by the.

Notes and references

(20) Sun, J.; Zuckermann, R. N. ACS Nano 2013, 7, 4715.
Peptide-peptoid hybrids

Peptoid units via Ugi:
- Natural amino acids
- Functional side chains
- Sequence specific

Globular structures